SEQUENCE LISTING

<110> XU, Ming-Qun
EVANS, Thomas C.
PRADHAN, Sriharsa
COMB, Donald G.
PAULUS, Henry
SUN, Luo
CHEN, Lixin
GHOSH, Inca
NEW ENGLAND BIOLABS, INC.
BOSTON BIOMEDICAL RESEARCH INSTITUTE

<120> METHOD FOR GENERATING SPLIT, NON-TRANSFERABLE GENES THAT ARE ABLE TO EXPRESS AN ACTIVE PROTEIN PRODUCT

<130> NEB-163-PCT

<140>
<141>

<150> 60/135,677
<151> 1999-05-24

<160> 134

<170> PatentIn Ver. 2.0

<210> 1
<211> 19
<212> DNA
<213> Escherichia coli

<400> 1
ggacggggaa ctaactatg 19

<210> 2
<211> 20
<212> DNA
<213> Escherichia coli

<400> 2
ccacgatgac gcaccacgcg 20

<210> 3
<211> 30
<212> DNA
<213> Escherichia coli

<400> 3
ggagggggca tatgaatggc gcacagtggg 30

<210> 4
<211> 25
<212> DNA
<213> Escherichia coli

<400> 4
ggggggtcat gataatttct ccaac 25

<210> 5
<211> 28
<212> DNA
<213> Escherichia coli

<400> 5
ccgggtggcg taattatgcc ggtttacg 28

<210> 6
<211> 28
<212> DNA
<213> Escherichia coli

<400> 6
cgtaaaccgg cataattacg ccacccgg 28

<210> 7
<211> 14
<212> PRT
<213> Synechocystis PCC6803

<400> 7
Leu Glu Lys Phe Ala Glu Tyr Cys Phe Asn Lys Ser Thr Gly
1 5 10

<210> 8
<211> 21
<212> PRT
<213> Escherichia coli

<400> 8
Cys Ala Gln Trp Val Val His Ala Leu Arg Ala Gln Gly Val Asn Thr
1 5 10 15

Val Phe Gly Tyr Gly

20

<210> 9
<211> 20
<212> PRT
<213> Escherichia coli

<400> 9
Cys Val Trp Pro Leu Val Pro Pro Gly Ala Ser Asn Ser Glu Met Leu
1 5 10 15

Glu Lys Leu Ser
20

<210> 10
<211> 26
<212> DNA
<213> Escherichia coli

<400> 10
gggggtcatg aatggcgcac agtggg 26

<210> 11
<211> 34
<212> DNA
<213> Escherichia coli

<400> 11
gcgcgctcga gttgatttaa cggctgctgt aatg 34

<210> 12
<211> 32
<212> DNA
<213> Escherichia coli

<400> 12
gcgcgaccgg ttgtgactgg cagcaacact gc 32

<210> 13
<211> 31
<212> DNA
<213> Escherichia coli

<400> 13
ggggggctgc agtcatgata atttctccaa c 31

<210> 14
<211> 22
<212> DNA
<213> MAIZE

<400> 14
atcagtacac agtcctgcca tc 22

<210> 15
<211> 20
<212> DNA
<213> MAIZE

<400> 15
gagacagccg ccgcaaccat 20

<210> 16
<211> 29
<212> DNA
<213> MAIZE

<400> 16
gggcccatat ggccaccgcc gccgccgcg 29

<210> 17
<211> 29
<212> DNA
<213> MAIZE

<400> 17
gggccctcga ggcttccttc aagaagagc 29

<210> 18
<211> 29
<212> DNA
<213> MAIZE

<400> 18
gggccaccgg tacatcaaag aagagcttg 29

<210> 19
<211> 31
<212> DNA
<213> MAIZE

<400> 19
ggggctgcat tcagtacaca gtcctgccat c 31

<210> 20
<211> 7
<212> PRT
<213> Synechocystis PCC6803

<400> 20
Leu Glu Lys Phe Ala Glu Tyr
1 5

<210> 21
<211> 7
<212> PRT
<213> Synechocystis PCC6803

<400> 21
Cys Phe Asn Lys Ser Thr Gly
1 5

<210> 22
<211> 21
<212> PRT
<213> MAIZE

<400> 22
Cys Lys Gly Ala Asp Ile Leu Val Glu Ser Leu Glu Arg Cys Gly Val
1 5 10 15

Arg Asp Val Phe Ala
20

<210> 23
<211> 21
<212> PRT
<213> MAIZE

<400> 23
Cys Ile Pro Ser Gly Gly Ala Phe Lys Asp Met Ile Leu Asp Gly Asp
1 5 10 15

Gly Arg Thr Val Tyr
20

<210> 24
<211> 44

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic, based on Salmonella typhimurium

<400> 24
ggatcctaag aaggagatat acccatggaa tccctgacgt taca 44

<210> 25
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic, based on Salmonella typhimurium

<400> 25
gtcgacgctc tcctgcagtt aggcaggcgt actcattc 38

<210> 26
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic, based on Salmonella typhimurium

<400> 26
gctttgctcc tggcggcttt accttgtggt aaaaccgc 38

<210> 27
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic, based on Salmonella typhimurium

<400> 27
gcggttttac cacaaggtaa agccgccagg agcaaagc 38

<210> 28
<211> 25

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 28
gcccctaaag acacaattat tcgcg 25

<210> 29
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 29
cagcggcgcc gtcatcagca gagcg 25

<210> 30
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 30
gcgaaccacc actaccaaca atttg 25

<210> 31
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 31
tatctccacg ccaaaggttt tcatt 25

<210> 32
<211> 21

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 32
gaatattgcc tgtcttttgg t 21

<210> 33
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 33
gttaaagcag ttagcagcga t 21

<210> 34
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 34
tgctgaatat tgcctgtctt ttgg 24

<210> 35
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 35
ccgttaaagc agttagcagc gatagc 26

<210> 36
<211> 44

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 36
ggatcctaag aaggagatat acccatggaa tccctgacgt taca 44

<210> 37
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 37
gatatcctgc agttaacctg gagagtgata ctgttgacc 39

<210> 38
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 38
gatatcccat gggacgctat ctggtcgagg gcgatg 36

<210> 39
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic,
based on Salmonella typhimurium

<400> 39
gtcgacgctc tcctgcagtt aggcaggcgt actcattc 38

<210> 40
<211> 31

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic from
Synechocystis species PCC6803

<400> 40
tgctgaatat gcgctgtctt ttggtaccga a 31

<210> 41
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic from
Synechocystis species PCC6803

<400> 41
ccgttaaacg ccgcagcagc gatagcgcc 29

<210> 42
<211> 178
<212> PRT
<213> Escherichia coli

<400> 42
Tyr Ala Val Asp Lys Ala Asp Leu Leu Leu Ala Leu Gly Val Arg Phe
1 5 10 15

Asp Asp Arg Val Thr Lys Ile Glu Ala Phe Ala Ser Arg Ala Lys Ile
20 25 30

Val His Val Asp Ile Asp Pro Ala Glu Ile Gly Lys Asn Lys Gln Pro
35 40 45

His Val Ser Ile Cys Ala Asp Val Lys Leu Ala Leu Gln Gly Met Asn
50 55 60

Ala Leu Leu Glu Gly Ser Thr Ser Lys Lys Ser Phe Asp Phe Gly Ser
65 70 75 80

Trp Asn Asp Glu Leu Asp Gln Gln Lys Arg Glu Phe Pro Leu Gly Tyr
85 90 95

Lys Thr Ser Asn Glu Glu Ile Gln Pro Gln Tyr Ala Ile Gln Val Leu
100 105 110

Asp Glu Leu Thr Lys Gly Glu Ala Ile Ile Gly Thr Gly Val Gly Gln
115 120 125

His Gln Met Trp Ala Ala Gln Tyr Tyr Thr Tyr Lys Arg Pro Arg Gln
130 135 140

Trp Leu Ser Ser Ala Gly Leu Gly Ala Met Gly Phe Gly Leu Pro Ala
145 150 155 160

Ala Ala Gly Ala Ser Val Ala Asn Pro Gly Val Thr Val Val Asp Ile
165 170 175

Asp Gly

<210> 43
<211> 179
<212> PRT
<213> Escherichia coli

<400> 43
Tyr Ala Val Asp Ser Ser Asp Leu Leu Leu Ala Phe Gly Val Arg Phe
1 5 10 15

Asp Asp Arg Val Thr Gly Lys Leu Glu Ala Phe Ala Ser Arg Ala Lys
20 25 30

Ile Val His Ile Asp Ile Asp Ser Ala Glu Ile Gly Lys Asn Lys Gln
35 40 45

Pro His Val Ser Ile Cys Ala Asp Ile Lys Leu Ala Leu Gln Gly Leu
50 55 60

Asn Ser Ile Leu Glu Ser Lys Glu Gly Lys Leu Lys Leu Asp Phe Ser
65 70 75 80

Ala Trp Arg Gln Glu Leu Thr Glu Gln Lys Val Lys His Pro Leu Asn
85 90 95

Phe Lys Thr Phe Gly Asp Ala Ile Pro Pro Gln Tyr Ala Ile Gln Val
100 105 110

Leu Asp Glu Leu Thr Asn Gly Asn Ala Ile Ile Ser Thr Gly Val Gly
115 120 125

Gln His Gln Met Trp Ala Ala Gln Tyr Tyr Lys Tyr Arg Lys Pro Arg

130 135 140

Gln Trp Leu Thr Ser Gly Gly Leu Gly Ala Met Gly Phe Gly Leu Pro
145 150 155 160

Ala Ala Ile Gly Ala Ala Val Gly Arg Pro Asp Glu Val Val Val Asp
165 170 175

Ile Asp Gly

<210> 44
<211> 179
<212> PRT
<213> Escherichia coli

<400> 44
Tyr Ala Val Asp Ser Ser Asp Leu Leu Leu Ala Phe Gly Val Arg Phe
1 5 10 15

Asp Asp Arg Val Thr Gly Lys Leu Glu Ala Phe Ala Ser Arg Ala Lys
20 25 30

Ile Val His Ile Asp Ile Asp Ser Ala Glu Ile Gly Lys Asn Lys Gln
35 40 45

Pro His Val Ser Ile Cys Ala Asp Ile Lys Leu Ala Leu Gln Gly Leu
50 55 60

Asn Ser Ile Leu Glu Ser Lys Glu Gly Lys Leu Lys Leu Asp Phe Ser
65 70 75 80

Ala Trp Arg Gln Glu Leu Thr Val Gln Lys Val Lys Tyr Pro Leu Asn
85 90 95

Phe Lys Thr Phe Gly Asp Ala Ile Pro Pro Gln Tyr Ala Ile Gln Val
100 105 110

Leu Asp Glu Leu Thr Asn Gly Ser Ala Ile Ile Ser Thr Gly Val Gly
115 120 125

Gln His Gln Met Trp Ala Ala Gln Tyr Tyr Lys Tyr Arg Lys Pro Arg
130 135 140

Gln Trp Leu Thr Ser Gly Gly Leu Gly Ala Met Gly Phe Gly Leu Pro
145 150 155 160

Ala Ala Ile Gly Ala Ala Val Gly Arg Pro Asp Glu Val Val Val Asp
165 170 175

Ile Asp Gly

<210> 45
<211> 180
<212> PRT
<213> Escherichia coli

<400> 45
Met Thr Met His Asn Ala Asp Val Ile Phe Ala Val Gly Val Arg Phe
1 5 10 15

Asp Asp Arg Thr Thr Asn Asn Leu Ala Lys Tyr Cys Pro Asn Ala Thr
20 25 30

Val Leu His Ile Asp Ile Asp Pro Thr Ser Ile Ser Lys Thr Val Thr
35 40 45

Ala Asp Ile Pro Ile Val Gly Asp Ala Arg Gln Val Leu Glu Gln Met
50 55 60

Leu Glu Leu Leu Ser Gln Glu Ser Ala His Gln Pro Leu Asp Glu Ile
65 70 75 80

Arg Asp Trp Trp Gln Gln Ile Glu Gln Trp Arg Ala Arg Gln Cys Leu
85 90 95

Lys Tyr Asp Thr His Ser Glu Lys Ile Lys Pro Gln Ala Val Ile Glu
100 105 110

Thr Leu Trp Arg Leu Thr Lys Gly Asp Ala Tyr Val Thr Ser Asp Val
115 120 125

Gly Gln His Gln Met Phe Ala Ala Leu Tyr Tyr Pro Phe Asp Lys Pro
130 135 140

Arg Arg Trp Ile Asn Ser Gly Gly Leu Gly Thr Met Gly Phe Gly Leu
145 150 155 160

Pro Ala Ala Leu Gly Val Lys Met Ala Leu Pro Glu Glu Thr Val Val
165 170 175

Cys Val Thr Gly
180

<210> 46
<211> 170
<212> PRT
<213> Escherichia coli

<400> 46
Phe Ala Val Gln Glu Cys Asp Leu Leu Ile Ala Val Gly Ala Arg Phe
1 5 10 15

Asp Asp Arg Val Thr Gly Lys Leu Asn Thr Ser Ala Pro His Ala Ser
20 25 30

Val Ile His Met Asp Ile Asp Pro Ala Glu Met Asn Lys Leu Arg Gln
35 40 45

Ala His Val Ala Leu Gln Gly Asp Leu Asn Ala Leu Leu Pro Ala Leu
50 55 60

Gln Gln Pro Leu Asn Gln Cys Asp Trp Gln Gln His Cys Ala Gln Leu
65 70 75 80

Arg Asp Glu His Ser Trp Arg Tyr Asp His Pro Gly Asp Ala Ile Tyr
85 90 95

Ala Pro Leu Leu Leu Lys Gln Leu Ser Asp Arg Lys Pro Ala Asp Cys
100 105 110

Val Val Thr Thr Asp Val Gly Gln His Gln Met Trp Ala Ala Gln His
115 120 125

Ile Ala His Thr Arg Pro Glu Asn Phe Ile Thr Ser Ser Gly Leu Gly
130 135 140

Thr Met Gly Phe Gly Leu Pro Ala Ala Val Gly Ala Gln Val Ala Arg
145 150 155 160

Pro Asn Asp Thr Val Val Cys Ile Ser Gly
165 170

<210> 47
<211> 35
<212> DNA
<213> Escherichia coli

<400> 47

gccttaatta accatgaggg aagcggtgat cgccg 35

<210> 48
<211> 34
<212> DNA
<213> Escherichia coli

<400> 48
tgcggtcgac tttgccgact accttggtga tctc 34

<210> 49
<211> 41
<212> DNA
<213> Escherichia coli

<400> 49
cccaagcttg gcgccatgag taaaggagaa gaacttttca c 41

<210> 50
<211> 36
<212> DNA
<213> Escherichia coli

<400> 50
gcgaccggtt tatttgtata gttcatccat gccatg 36

<210> 51
<211> 39
<212> DNA
<213> Escherichia coli

<400> 51
agggaattcg tcgacaaatt tgctgaatat tgcctgtct 39

<210> 52
<211> 38
<212> DNA
<213> Escherichia coli

<400> 52
ggcctcgagt tatttaattg tcccagcgtc aagtaatg 38

<210> 53
<211> 41
<212> DNA
<213> Escherichia coli

<400> 53

agctttgttt aaaccatggt taaagttatc ggtcgtagat c 41

<210> 54
<211> 43
<212> DNA
<213> Escherichia coli

<400> 54
cagcgtcgac ggcgccgtgg gatttgttaa agcagttagc agc 43

<210> 55
<211> 31
<212> DNA
<213> Escherichia coli

<400> 55
catgccatgg gggaagcggt gatcgccgaa g 31

<210> 56
<211> 39
<212> DNA
<213> Escherichia coli

<400> 56
acgcgagctc ttatttaatt gtcccagcgt caagtaatg 39

<210> 57
<211> 34
<212> DNA
<213> Escherichia coli

<400> 57
cgaattctat ggttaaagtt atcggtcgta gatc 34

<210> 58
<211> 36
<212> DNA
<213> Escherichia coli

<400> 58
agcccgcggt tatttgtata gttcatccat gccatg 36

<210> 59
<211> 154
<212> DNA
<213> Nicotiana tabacum

<400> 59

gaatagatct acatacacct tggttgacac gagtatataa gtcatgttat actgttgaat 60
aacaagcctt ccattttcta ttttgatttg tagaaaacta gtgtgcttgg gagtccctga 120
tgattaaata aaccaagatt ttaccttaat taag 154

<210> 60
<211> 151
<212> DNA
<213> Nicotiana tabacum

<400> 60
gatcctggcc tagtctatag gaggttttga aaagaaagga gcaataatca ttttcttgtt 60
ctatcaagag ggtgctattg ctcctttctt tttttctttt tatttattta ctagtatttt 120
acttacatag actttttgt ttacgtattc t 151

<210> 61
<211> 185
<212> DNA
<213> Nicotiana tabacum

<400> 61
catatggcgt ccatgatctc ctcgtccgcg gtgaccacgg tcagccgcgc gtccacggtg 60
cagtcggccg cggtggcccc gttcggcggc ctcaagtcca tgaccggctt cccggtcaag 120
aaggtcaaca cggacatcac gtccatcacg agcaacggcg gcagggtgaa gtgcatgcga 180
agagc 185

<210> 62
<211> 6232
<212> DNA
<213> Unknown

<220>
<223> nucleotides 1-2492: E. coli vector pLITMUS28 (New
England Biolabs, Inc.)

<220>
<223> nucleotides 2493-5993: Nicotiana tabaceum

<220>
<223> Nucleotides 5993-6232: E.coli vector pLITMUS28
(New England Biolabs, Inc.)

<400> 62
gttaactacg tcaggtggca cttttcgggg aaatgtgcgc ggaacccta tttgtttatt 60
tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat aaatgcttca 120
ataatattga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc ttattccctt 180
ttttgcggca ttttgccttc ctgtttttgc tcacccagaa acgctggtga aagtaaaaga 240
tgctgaagat cagttgggtg cacgagtggg ttacatcgaa ctggatctca acagcggtaa 300
gatccttgag agttttcgcc ccgaagaacg ttctccaatg atgagcactt ttaaagttct 360

gctatgtggc gcggtattat cccgtgttga cgccgggcaa gagcaactcg gtcgccgcat 420
acactattct cagaatgact tggttgagta ctcaccagtc acagaaaagc atcttacgga 480
tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata acactgcggc 540
caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt tgcacaacat 600
gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag ccataccaaa 660
cgacgagcgt gacaccacga tgcctgtagc aatggcaaca acgttgcgca aactattaac 720
tggcgaacta cttactctag cttcccggca acaattaata gactggatgg aggcggataa 780
agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg ctgataaatc 840
tggagccggt gagcgtgggt ctcgcggtat cattgcagca ctggggccag atggtaagcc 900
ctcccgtatc gtagttatct acacgacggg gagtcaggca actatggatg aacgaaatag 960
acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag accaagttta 1020
ctcatatata ctttagattg atttaccccg gttgataatc agaaaagccc caaaaacagg 1080
aagattgtat aagcaaatat ttaaattgta aacgttaata ttttgttaaa attcgcgtta 1140
aatttttgtt aaatcagctc attttttaac caataggccg aaatcggcaa aatcccttat 1200
aaatcaaaag aatagcccga gatagggttg agtgttgttc cagtttggaa caagagtcca 1260
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 1320
ccactacgtg aaccatcacc caaatcaagt tttttggggt cgaggtgccg taaagcacta 1380
aatcggaacc ctaaagggag cccccgattt agagcttgac ggggaaagcg aacgtggcga 1440
gaaaggaagg gaagaaagcg aaaggagcgg gcgctagggc gctggcaagt gtagcggtca 1500
cgctgcgcgt aaccaccaca cccgccgcgc ttaatgcgcc gctacagggc gcgtaaaagg 1560
atctaggtga agatcctttt tgataatctc atgaccaaaa tcccttaacg tgagttttcg 1620
ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt 1680
ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggt ggtttgtttg 1740
ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata 1800
ccaaatactg ttcttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca 1860
ccgcctacat acctcgctct gctaatcctg ttaccagtgg ctgctgccag tggcgataag 1920
tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc 1980
tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga 2040
tacctacagc gtgagctatg agaaagcgcc acgcttcccg aagggagaaa ggcggacagg 2100
tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc agggggaaac 2160
gcctggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg 2220
tgatgctcgt caggggggcg gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg 2280
ttcctggcct tttgctggcc ttttgctcac atgtaatgtg agttagctca ctcattaggc 2340
accccaggct ttacacttta tgcttccggc tcgtatgttg tgtggaattg tgagcggata 2400
acaatttcac acaggaaaca gctatgacca tgattacgcc aagctacgta atacgactca 2460
ctagtgggca gatcttcgaa tgcatcgcgc gcttgacgat atagcaattt tgcttggatt 2520
tatcagtcga agcaggagac aatatacctt gatattctcg atcattcttt gattcaaagc 2580
atcgttccat ctcaattgaa aaagcaaata acgtttcaag aacaaatcta gttctgcttc 2640
cgtgttgctt ttgtattgtt ttttcttttt acccttcttt gtgtctgatt ccgcgtaatc 2700
tttttttaaga gcgttttgat gttttgagag aacagggccc agatttcctt tgttttctat 2760
atctgatcca cgctcttttt ctccttgact tgcgggttct tttgcttctt gaattcgatt 2820
ctttattttt ttatttgatc gtagaaaaaa gttttgtttt tggtttttat tgatgttttt 2880
attttgacta acattttcat ttgtattcaa atttaaaaga agtaatttgc ttggtataat 2940
ccacggtttt attttatata cattataaag tggtacaaat tctgggaaga accaaaattc 3000
cagattcaat atgggacgat ttaatatttt ttcattcatt cccatccaat caaaaaaggc 3060
ttttttcgaa tttttttgat tgttttctgg attttgatga atcgtaagat aaaaaaagcc 3120
tttttatca attttatcaa ttatttgata attattaata ccaattttag tatttggatt 3180
actgttggta tcgatcttaa cccaggcctc aatatcttct tttgtctaa gagaaaaatg 3240

gataattttc caatcaaaat attttctatc gagatttctt tctatatata gaatattgcc 3300
ttttcttaga taattattga tatgaagatt gccgagcata tcaaaaaggt tgtgtttgga 3360
cgtgttggaa ttagaagaaa tttcgaggtt cttatttact tgaaagggta atctagaaat 3420
aaaagagtca tttttttttt cataattaat cgatttatat gctaaaagat catatctata 3480
acattttga aaattatctt tttggtttgc taatgaatag agctcagaat catttcttt 3540
tttgtaatga attaattggt ctttttcata tgaattccat ttgtttaaat ttcgattttg 3600
agccatacaa ccttgattaa ccctatttcg ccatttttgt ggcattaatc tagaccatct 3660
aatctgagat aaatcgtacg agaatactca atcatgaata aatgcaagaa aataacctct 3720
ccttcttttt ctataatgta aacaaaaaag tctatgtaag taaaatacta gtaaataaat 3780
aaaaagaaaa aaagaaagga gcaatagcac cctcttgata gaacaagaaa atgattattg 3840
ctcctttctt ttcaaaacct cctatagact aggccaggat cctcgagctt aattaaggta 3900
aaatcttggt ttatttaatc atcagggact cccaagcaca ctagtttct acaaatcaaa 3960
atagaaaata gaaaatggaa ggctttttat tcaacagtat aacatgactt atatactcgt 4020
gtcaaccaag gtgtatgtag atctattcct gcaggatatc tggatccacg aagcttccca 4080
tgggaataga tctacataca ccttggttga cacgagtata taagtcatgt tatactgttg 4140
aataaaaagc cttccatttt ctattttgat ttgtagaaaa ctagtgtgct tgggagtccc 4200
tgatgattaa ataaaccaag attttaccgt ttaaacaccg gtgatcctgg cctagtctat 4260
aggaggtttt gaaaagaaag gagcaataat catttcttg ttctatcaag agggtgctat 4320
tgctcctttc tttttttctt tttatttatt tactagtatt ttacttacat agactttttt 4380
gtttacatta tagaaaaaga aggagaggtt atttcttgc atttattcat gattgagtat 4440
tctcctaggc gtattgataa tgccgtctta accagttttt ccattgattg attctataac 4500
tctgaagttt cttatgtttt aattcagaat gaaatattcc tagtgttcga aaatagtcct 4560
ttattttagt cttaaggaaa aaagacgttc tgttatattg aagaacagat cttaatttag 4620
acaaattaat aacttggggt tgtgataatt tgtaaaatac atatgcttgt gataagtagg 4680
ataaatcaaa aaaaatatgt gaatttttct tactaatatt ataaagtgac tttttatag 4740
tcgaaataaa gtgaattttt ttttgattat taattttttc ttgatttatt tcattattgg 4800
aaatgtattt atcaatcaat ttgtttgttg attcaagaaa gagttgtgta ttaattctgg 4860
gaatattaat gatagataaa aatagatcga tgtataatct ttgaatgaat aattttagaa 4920
aataatggaa tttccatatt aatcgagtat ttcttctttt taatatttgg aaaatctttt 4980
ttggcgattc gaatttttta atattatttg ttttattagg actaatgtct atttctggag 5040
ttactttctt tttctctttt gtaattcttt ctatttgatt tttgattgta cttgttctat 5100
cagtcaaatc cttcattttg ctttctatca gtgaagaatt tggccaattt ccagattcaa 5160
tttgactaaa tgattcgtta attatctgat tactcattag agaatctttt tctttttcg 5220
tttcattcga ttcatctatt tctttgagtc taaataatac aattggattt acttttgaaa 5280
gttctttttt catttttttt ataaatagac tacttttgat aagccatttt ttggtttctt 5340
ttgaaattct tcgaaataat tttattttc ctttgaaaac ttttagagtt ataaaatatt 5400
tctttttgaa ttttccaatt tttttttcga gttccttaaa aatgggctca aaaaaagaag 5460
ggcgttttcg gggagaacca aagggaagtt cagcttccat tccccaaact gttaaaaaac 5520
aaaaatcatc tttttgtttt ttctttttca ttagctctcc acgggaggag tacagtttag 5580
atatatgcca aggtttcaga caaaaaggaa ataatatttt gatctgaatg ccatctttca 5640
accaattttt tggaaattct gtttctgata attgaacacc attataagta catttaatat 5700
gcatttctct attccattcc tgcaaatctt cagaccattc aggaagttgc aagactaaca 5760
tacgcccgag atttttggct attatcaatg aaggtaatac aatatatttt cgaagaattg 5820
attgagttat taacatgtaa cctcttatta tttgcgcaaa aggaatggta tcccaggctt 5880
ctgctatctc tatccgtgct ttttcctttc ttttgttctc cccttttttg tccttttcct 5940
ttttctcttc tctttttgtt tgttcttctc tagactctag aatcttgaat tcggtaccct 6000
ctagtcaagg ccttaagtga gtcgtattac ggactggccg tcgttttaca acgtcgtgac 6060
tgggaaaacc ctggcgttac ccaacttaat cgccttgcag cacatccccc tttcgccagc 6120

tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc aacagttgcg cagcctgaat 6180
ggcgaatggc gcttcgcttg gtaataaagc ccgcttcggc gggctttttt tt 6232

<210> 63
<211> 6477
<212> DNA
<213> Unknown

<220>
<223> Nucleotides 1-2482: E. coli vector pLITMUS28 (New England Biolabs, Inc.)

<220>
<223> Nucleotides 2493-6242: Nicotiana tabaceum

<220>
<223> Nucleotides 6243-6477: E. coli vector pLITMUS28 (New England Biolabs, Inc.)

<400> 63
gttaactacg tcaggtggca cttttcgggg aaatgtgcgc ggaaccccta tttgtttatt 60
tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat aaatgcttca 120
ataatattga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc ttattccctt 180
ttttgcggca ttttgccttc ctgtttttgc tcacccagaa acgctggtga aagtaaaaga 240
tgctgaagat cagttgggtg cacgagtggg ttacatcgaa ctggatctca acagcggtaa 300
gatccttgag agttttcgcc ccgaagaacg ttctccaatg atgagcactt ttaaagttct 360
gctatgtggc gcggtattat cccgtgttga cgccgggcaa gagcaactcg gtcgccgcat 420
acactattct cagaatgact tggttgagta ctcaccagtc acagaaaagc atcttacgga 480
tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata acactgcggc 540
caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt tgcacaacat 600
gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag ccataccaaa 660
cgacgagcgt gacaccacga tgcctgtagc aatggcaaca acgttgcgca aactattaac 720
tggcgaacta cttactctag cttcccggca acaattaata gactggatgg aggcggataa 780
agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg ctgataaatc 840
tggagccggt gagcgtgggt ctcgcggtat cattgcagca ctggggccag atggtaagcc 900
ctcccgtatc gtagttatct acacgacggg gagtcaggca actatggatg aacgaaatag 960
acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag accaagttta 1020
ctcatatata ctttagattg atttaccccg gttgataatc agaaaagccc caaaaacagg 1080
aagattgtat aagcaaatat ttaaattgta aacgttaata ttttgttaaa attcgcgtta 1140
aatttttgtt aaatcagctc attttttaac caataggccg aaatcggcaa aatcccttat 1200
aaatcaaaag aatagcccga gatagggttg agtgttgttc cagtttggaa caagagtcca 1260
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 1320
ccactacgtg aaccatcacc caaatcaagt tttttggggt cgaggtgccg taaagcacta 1380
aatcggaacc ctaaagggag cccccgattt agagcttgac ggggaaagcg aacgtggcga 1440
gaaaggaagg gaagaaagcg aaaggagcgg gcgctagggc gctggcaagt gtagcggtca 1500
cgctgcgcgt aaccaccaca cccgccgcgc ttaatgcgcc gctacagggc gcgtaaaagg 1560
atctaggtga agatcctttt tgataatctc atgaccaaaa tcccttaacg tgagttttcg 1620
ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt 1680

ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggt ggtttgtttg 1740
ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata 1800
ccaaatactg ttcttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca 1860
ccgcctacat acctcgctct gctaatcctg ttaccagtgg ctgctgccag tggcgataag 1920
tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc 1980
tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga 2040
tacctacagc gtgagctatg agaaagcgcc acgcttcccg aagggagaaa ggcggacagg 2100
tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc agggggaaac 2160
gcctggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg 2220
tgatgctcgt caggggggcg gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg 2280
ttcctggcct tttgctggcc ttttgctcac atgtaatgtg agttagctca ctcattaggc 2340
accccaggct ttacacttta tgcttccggc tcgtatgttg tgtggaattg tgagcggata 2400
acaatttcac acaggaaaca gctatgacca tgattacgcc aagctacgta atacgactca 2460
ctagtgggca gatcttcgaa tgcatcgcgc gcaattcacc gccgtatggc tgaccggcga 2520
ttactagcga ttccggcttc atgcaggcga gttgcagcct gcaatccgaa ctgaggacgg 2580
gtttttgggg ttagctcacc ctcgcgggat cgcgaccctt tgtcccggcc attgtagcac 2640
gtgtgtcgcc cagggcataa ggggcatgat gacttgacgt catcctcacc ttcctccggc 2700
ttatcaccgg cagtctgttc agggttccaa actcaacgat ggcaactaaa cacgagggtt 2760
gcgctcgttg cgggacttaa cccaacacct tacggcacga gctgacgaca gccatgcacc 2820
acctgtgtcc gcgttcccga aggcacccct ctctttcaag aggattcgcg gcatgtcaag 2880
ccctggtaag gttcttcgct ttgcatcgaa ttaaaccaca tgctccaccg cttgtgcggg 2940
cccccgtcaa ttcctttgag tttcattctt gcgaacgtac tccccaggcg ggatacttaa 3000
cgcgttagct acagcactgc acgggtcgat acgcacagcg cctagtatcc atcgtttacg 3060
gctaggacta ctggggtatc taatcccatt cgctccccta gctttcgtct ctcagtgtca 3120
gtgtcggccc agcagagtgc tttcgccgtt ggtgttcttt ccgatctcta cgcatttcac 3180
cgctccaccg gaaattccct ctgcccctac cgtactccag cttggtagtt tccaccgcct 3240
gtccagggtt gagccctggg atttgacggc ggacttaaaa agccacctac agacgcttta 3300
cgcccaatca ttccggataa cgcttgcatc ctctgtatta ccgcggctgc tggcacagag 3360
ttagccgatg cttattcccc agataccgtc attgcttctt ctccgggaaa agaagttcac 3420
gacccgtggg ccttctacct ccacgcggca ttgctccgtc agctttcgcc cattgcggaa 3480
aattccccac tgctgcctcc cgtaggagtc tgggccgtgt ctcagtccca gtgtggctga 3540
tcatcctctc ggaccagcta ctgatcatcg ccttggtaag ctattgcctc accaactagc 3600
taatcagacg cgagcccctc ctcgggcgga ttcctccttt tgctcctcag cctacggggt 3660
attagcagcc gtttccagct gttgttcccc tcccaagggc aggttcttac gcgttactca 3720
cccgtccgcc actggaaaca ccacttcccg tccgacttgc atgtgttaag catgccgcca 3780
gcgttcatcc tgagccagga tcgaactctc catgagattc atagttgcat tacttatagc 3840
ttccttgttc gtagacaaag cggattcgga attgtctttc attccaaggc ataacttgta 3900
tccatgcgct tcatattcgc ccggagttcg ctcccagaaa tatagccatc cctgcccccct 3960
cacgtcaatc ccacgagcct cttatccatt ctcattgaac gacggcgggg gagcaaatcc 4020
aactagaaaa actcacattg ggcttaggga taatcaggct cgaactgatg acttccacca 4080
cgtcaaggtg acactctacc gctgagttat atcccttccc cgccccatcg agaaatagaa 4140
ctgactaatc ctaagtcaaa ggcgtacgag aatactcaat catgaataaa tgcaagaaaa 4200
taacctctcc ttcttttttct ataatgtaaa caaaaaagtc tatgtaagta aaatactagt 4260
aaataaataa aaagaaaaaa agaaaggagc aatagcaccc tcttgataga acaagaaaat 4320
gattattgct cctttctttt caaaacctcc tatagactag gccaggatcc tcgagcttaa 4380
ttaaggtaaa atcttggttt atttaatcat cagggactcc caagcacact agtttctac 4440
aaatcaaaat agaaaataga aaatggaagg ctttttattc aacagtataa catgacttat 4500
atactcgtgt caaccaaggt gtatgtagat ctattcctgc aggatatctg gatccacgaa 4560

gcttcccatg ggaatagatc tacatacacc ttggttgaca cgagtatata agtcatgtta 4620
tactgttgaa taaaaagcct tccattttct attttgattt gtagaaaact agtgtgcttg 4680
ggagtccctg atgattaaat aaaccaagat tttaccgttt aaacaccggt gatcctggcc 4740
tagtctatag gaggttttga aaagaaagga gcaataatca ttttcttgtt ctatcaagag 4800
ggtgctattg ctcctttctt tttttctttt tatttattta ctagtatttt acttacatag 4860
actttttgt ttacattata gaaaaagaag gagaggttat tttcttgcat ttattcatga 4920
ttgagtattc tcctagggtc gagaaactca acgccactat tcttgaacaa cttggagccg 4980
ggccttcttt tcgcactatt acggatatga aaataatggt caaaatcgga ttcaattgtc 5040
aactgcccct atcggaaata ggattgacta ccgattccga aggaactgga gttacatctc 5100
ttttccattc aagagttctt atgcgtttcc acgccccttt gagaccccga aaaatggaca 5160
aattcctttt cttaggaaca catacaagat tcgtcactac aaaaaggata atggtaaccc 5220
taccattaac tacttcattt atgaatttca tagtaataga aatacatgtc ctaccgagac 5280
agaatttgga acttgctatc ctcttgccta gcaggcaaag atttacctcc gtggaaagga 5340
tgattcattc ggatcgacat gagagtccaa ctacattgcc agaatccatg ttgtatattt 5400
gaaagaggtt gacctccttg cttctctcat ggtacactcc tcttcccgcc gagccccttt 5460
tctcctcggt ccacagagac aaaatgtagg actggtgcca acaattcatc agactcacta 5520
agtcgggatc actaactaat actaatctaa tataatagtc taatatatct aatataatag 5580
aaaatactaa tataatagaa aagaactgtc ttttctgtat actttccccg gttccgttgc 5640
taccgcgggc tttacgcaat cgatcggatt agatagatat cccttcaaca taggtcatcg 5700
aaaggatctc ggagacccac caaagtacga aagccaggat ctttcagaaa acggattcct 5760
attcaaagag tgcataaccg catggataag ctcacactaa cccgtcaatt tgggatccaa 5820
attcgagatt ttccttggga ggtatcggga aggatttgga atggaataat atcgattcat 5880
acagaagaaa aggttctcta ttgattcaaa cactgtacct aacctatggg atagggatcg 5940
aggaagggga aaaaccgaag atttcacatg gtacttttat caatctgatt tatttcgtac 6000
ctttcgttca atgagaaaat gggtcaaatt ctacaggatc aaacctatgg gacttaagga 6060
atgatataaa aaaaagagag ggaaaatatt catattaaat aaatatgaag tagaagaacc 6120
cagattccaa atgaacaaat tcaaacttga aaaggatctt ccttattctt gaagaatgag 6180
gggcaaaggg attgatcaag aaagatcttt tgttcttctt atatataaga tcgtgatggt 6240
accctctagt caaggcctta agtgagtcgt attacggact ggccgtcgtt ttacaacgtc 6300
gtgactggga aaaccctggc gttacccaac ttaatcgcct tgcagcacat ccccctttcg 6360
ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttcccaacag ttgcgcagcc 6420
tgaatggcga atggcgcttc gcttggtaat aaagcccgct tcggcgggct ttttttt 6477

<210> 64
<211> 31
<212> DNA
<213> Nicotiana tabacum

<400> 64
aactgcagga atagatctac atacaccttg g 31

<210> 65
<211> 42
<212> DNA
<213> Nicotiana tabacum

<400> 65
ccgctcgagc ttaattaagg taaaatcttg gtttatttaa tc 42

<210> 66
<211> 33
<212> DNA
<213> Nicotiana tabacum

<400> 66
gcgaccggtg atcctggcct agtctatagg agg 33

<210> 67
<211> 34
<212> DNA
<213> Nicotiana tabacum

<400> 67
aggcctagga gaatactcaa tcatgaataa atgc 34

<210> 68
<211> 34
<212> DNA
<213> Nicotiana tabacum

<400> 68
ttggcgcgct tgacgatata gcaattttgc ttgg 34

<210> 69
<211> 34
<212> DNA
<213> Nicotiana tabacum

<400> 69
ttgcgtacga tttatctcag attagatggt ctag 34

<210> 70
<211> 35
<212> DNA
<213> Nicotiana tabacum

<400> 70
ttgcctaggc gtattgataa tgccgtctta accag 35

<210> 71
<211> 34
<212> DNA
<213> Nicotiana tabacum

<400> 71
aggggtaccg aattcaagat tctagagtct agag 34

<210> 72
<211> 34
<212> DNA
<213> Nicotiana tabacum

<400> 72
ttggcgcgca attcaccgcc gtatggctga ccgg 34

<210> 73
<211> 34
<212> DNA
<213> Nicotiana tabacum

<400> 73
ttgcgtacgc ctttgactta ggattagtca gttc 34

<210> 74
<211> 34
<212> DNA
<213> Nicotiana tabacum

<400> 74
ttgcctaggg tcgagaaact caacgccact attc 34

<210> 75
<211> 35
<212> DNA
<213> Nicotiana tabacum

<400> 75
aggggtacca tcacgatctt atatataaga agaac 35

<210> 76
<211> 250
<212> DNA
<213> Nicotiana tabacum

<400> 76
gaattgtgag cgctcacaat tctaggatgt taattgcgcc gacatcataa cggttctggc 60
aaatattctg aaatgagctg ttgacaatta atcatcggct cgtataatgt gtggaattgt 120
gagcggataa caatttcaca caggaaacag accatggtga attctagagc tcgaggatcc 180
gcggtacccg ggcatgcatt cgaagcttcc ttaagcggcc gtcgaccgat gcccttgaga 240
gccttcaacc 250

<210> 77
<211> 5
<212> PRT

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 77
Cys Leu Asn Ile Gln
1 5

<210> 78
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 78
Val Phe Lys His Ala
1 5

<210> 79
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 79
Leu Phe Lys Gln Pro
1 5

<210> 80
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 80
Cys Leu Asn Ser Asp
1 5

<210> 81
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 81
Cys Leu Asn Ile Ser
1 5

<210> 82
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 82
Cys Leu Asn Thr Asp
1 5

<210> 83
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 83
Cys Leu Asn Asn Arg
1 5

<210> 84

<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 84
Cys Leu Asn Ser Cys
1 5

<210> 85
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 85
Cys Leu Asn Ser Asp
1 5

<210> 86
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 86
Cys Leu Asn Thr Leu
1 5

<210> 87
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the

ends of the Tn7 transposon

<400> 87
Val Phe Lys Gln Pro
1 5

<210> 88
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 88
Cys Leu Asn Ser Met
1 5

<210> 89
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 89
Cys Leu Asn Asn Tyr
1 5

<210> 90
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 90
Cys Leu Asn Met Ala
1 5

<210> 91
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 91
Val Phe Lys His Lys
1 5

<210> 92
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 92
Cys Leu Asn Thr Lys
1 5

<210> 93
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 93
Cys Leu Asn Lys Asp
1 5

<210> 94
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 94
Met Phe Lys Gln Ile
1 5

<210> 95
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 95
Cys Leu Asn Ile Ile
1 5

<210> 96
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 96
Leu Phe Lys His Glu
1 5

<210> 97
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 97
Val Phe Lys His Phe

1 5

<210> 98
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 98
Cys Leu Asn Ser Val
1 5

<210> 99
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 99
Val Phe Lys Gln Ile
1 5

<210> 100
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 100
Met Phe Lys Gln Ala
1 5

<210> 101
<211> 5
<212> PRT

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 101
Leu Phe Lys His His
1 5

<210> 102
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 102
Leu Phe Lys His Gln
1 5

<210> 103
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 103
Met Phe Lys His Val
1 5

<210> 104
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 104
Val Phe Lys Gln Lys
1 5

<210> 105
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 105
Leu Phe Lys Gln Gln
1 5

<210> 106
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 106
Leu Phe Lys His Ser
1 5

<210> 107
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 107
Cys Leu Asn Thr Gly
1 5

<210> 108

<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 108
Cys Leu Asn Ser Arg
1 5

<210> 109
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 109
Val Phe Lys His Leu
1 5

<210> 110
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 110
Cys Leu Asn Asn Ile
1 5

<210> 111
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the

ends of the Tn7 transposon

<400> 111
Leu Phe Lys His Gln
1 5

<210> 112
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 112
Cys Leu Asn Lys His
1 5

<210> 113
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 113
Met Phe Lys Gln Tyr
1 5

<210> 114
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 114
Cys Leu Asn Lys Gln
1 5

<210> 115
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 115
Cys Leu Asn Met Ser
1 5

<210> 116
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 116
Leu Cys Leu Asn Ile Leu Ala
1 5

<210> 117
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 117
Asn Cys Leu Asn Ile Asn Ala
1 5

<210> 118
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 118
Leu Met Phe Lys His Leu Ser
1 5

<210> 119
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 119
Thr Leu Phe Lys His Thr Arg
1 5

<210> 120
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 120
Lys Val Phe Lys Gln Lys Glu
1 5

<210> 121
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 121
His Leu Val Phe Lys His Leu

1 5

<210> 122
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 122
Leu Cys Leu Asn Thr Leu Leu
1 5

<210> 123
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 123
Leu Cys Leu Asn Asn Leu Val
1 5

<210> 124
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 124
Glu Val Phe Lys His Glu Gly
1 5

<210> 125
<211> 7
<212> PRT

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 125
Lys Val Phe Lys Gln Lys Gly
1 5

<210> 126
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 126
Thr Cys Leu Asn Thr Thr Ile
1 5

<210> 127
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 127
Met Cys Leu Asn Asn Met Asn
1 5

<210> 128
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the
ends of the Tn7 transposon

<400> 128
Leu Leu Phe Lys Gln Leu Arg
1 5

<210> 129
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 129
Arg Cys Leu Asn Asn Arg Leu
1 5

<210> 130
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 130
Met Val Phe Lys Gln Met Ala
1 5

<210> 131
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 131
Ala Met Phe Lys Gln Ala Thr
1 5

<210> 132

<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 132
Leu Val Phe Lys His Leu Asp
1 5

<210> 133
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 133
Lys Met Phe Lys Gln Lys Thr
1 5

<210> 134
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: based on the ends of the Tn7 transposon

<400> 134
Tyr Cys Leu Asn Asn Tyr Phe
1 5